DOCUMENT RESUME

ED 340 743 TM 017 809

AUTHOR Youngjohn, James R.; And Others
TITLE Test-Retest Reliability of Computerized, Everyday Memory Measures and Traditional Memory Tests.
PUB DATE Aug 91
NOTE 12p.; Paper presented at the Annual Meeting of the American Psychological Association (99th, San Francisco, CA, August 16-20, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adults; Comparative Testing; *Computer Assisted Testing; Computer Simulation; *Drills (Practice); *Memory; Neuropsychology; Psychological Testing; Psychometrics; *Test Reliability; Volunteers
IDENTIFIERS Benton Visual Retention Test; Memory Tasks; *Test Retest Reliability; Wechsler Adult Intelligence Scale; Wechsler Memory Scale

ABSTRACT

Test-retest reliabilities and practice effect magnitudes were considered for nine computer-simulated tasks of everyday cognition and five traditional neuropsychological tests. The nine simulated everyday memory tests were from the Memory Assessment Clinic battery as follows: (1) simple reaction time while driving; (2) divided attention (driving while listening to the radio); (3) misplaced objects; (4) name-face association; (5) recognition of faces--signal detection; (6) recognition of faces--delayed non-match to sample; (7) telephone dialing; (8) first-last names; and (9) grocery list selective reminding. The traditional tests (Logical Memory and Paired Associate Learning from the Wechsler Memory Scale, Digit Span and Digit Symbol from the Wechsler Adult Intelligence Scale, and the Benton Visual Retention Test) were administered. Subjects were 56 male and 59 female volunteers aged between 17 and 82 years. Most measures show significant practice effects on reevaluation. Test-retest reliabilities of the computerized recall measures are equal or superior to the traditional memory tests, but the traditional measures of attention and concentration had reliabilities superior to the everyday tests. Test-retest reliabilities of everyday recognition memory tests were disappointing. Advantages of computerization of psychometric assessment and the need for multiple test forms are discussed. Two tables present study data. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
James R. Youngjohn
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED340743

TEST-RETEST RELIABILITY OF COMPUTERIZED, EVERYDAY AND TRADITIONAL MEMORY TESTS

James R. Youngjohn, Glenn J. Larrabee, and Thomas H. Crook, III
Memory Assessment Clinics, Inc.

Poster presented at the annual convention of the American Psychological Association, San Francisco, 1991.

Address for reprints: James R. Youngjohn, Ph.D.
Memory Assessment Clinics, Inc.
Lincoln Plaza Medical Center
7125 East Lincoln Drive
Scottsdale, Arizona 85253

TM017809

BEST COPY AVAILABLE

TEST-RETEST RELIABILITY OF COMPUTERIZED, EVERYDAY MEMORY MEASURES AND TRADITIONAL MEMORY TESTS

James R. Youngjohn, Glenn J. Larrabee, and Thomas H. Crook, III
Memory Assessment Clinics, Inc.

Abstract

Test-retest reliabilities and practice effect magnitudes are considered for nine computer-simulated tasks of everyday cognition and five traditional neuropsychological tests. Most measures showed significant practice effects upon reevaluation. Test-retest reliabilities of the computerized recall measures were equal to or superior to the traditional memory tests, but traditional measures of attention and concentration had superior reliabilities to the everyday tests. Test-retest reliabilities of everyday recognition memory tests were disappointing. Advantages of computerization of psychometric assessment and the need for multiple test forms are discussed.

TEST-RETEST RELIABILITY OF COMPUTERIZED, EVERYDAY AND TRADITIONAL MEMORY TESTS

In the present investigation, we examined test-retest reliability for nine of the computerized, everyday memory tests comprising the Memory Assessment Clinic (MAC) battery (Crook & Larrabee, 1988; Crook, Salama, and Gobert, 1986; Larrabee & Crook, 1989a). We compared these findings to test-retest reliability for a number of traditional neuropsychological and memory tests. Of particular interest were the comparisons between the traditional and MAC tests sharing close structural resemblance. These include Telephone Dialing (an everyday memory analogue of Wechsler Adult Intelligence Scale Digit Span; Wechsler, 1955), First-Last Names and Name-Face Memory (everyday memory tests using the Wechsler Memory Scale Paired Associate Learning format; Wechsler, 1945), and Divided Attention Recall (recall of radio weather and traffic reports, scored in the same manner as Wechsler Memory Scale Logical Memory). Finally, we examined the magnitude of the practice effects across the two administrations for both the MAC and the traditional tests.

METHODS

Participants

One hundred and fifteen normal volunteers between the ages of 17 and 82 years of age (mean = 48.87 years, SD = 19.30) participated in the study. The group consisted of 56 males and 59 females.

Procedure

Retesting was scheduled 21 days after the initial exam (mean test-retest interval = 20.94 days, SD = 5.09, minimum = 13 days, maximum = 41 days). Descriptions of the MAC tests follow.

Simple Reaction Time: The task of driving a car and responding appropriately to changing traffic signals is simulated. Both lift and travel time are measured as subjects move back and forth from brake to accelerator

pedals in response to traffic light signals.

Divided Attention: This task is a variation of WMS Logical Memory. Subjects listen to simulated radio weather and traffic reports while performing the simulated driving task described above. Immediate recall is assessed in a similar manner to Logical Memory.

Misplaced Objects: This is a test of object location recall. Subjects are required to place 20 common objects into the schematic representation of a 12 room house and then recall their locations 40 minutes later. Two recall attempts are allowed.

Name-Face Association: Subjects are presented with live color video recordings (stored on laser disk) of 14 individuals introducing themselves by common first names. Recall is assessed by showing the same individuals stating the city where they live and asking the subject to provide the name of each person. There are three learning trials. Total name-face pairs recalled is summed over the learning trials. Delayed recall is assessed after 40 minutes.

Recognition of Faces-Signal Detection: In this test subjects are presented with 156 individual facial photographs. Signal detection methods are used to control for varying response criteria among subjects, and the statistic d' is calculated for each delay interval. Delayed recognition memory is also assessed by presenting 20 faces after 40 minutes, 10 of which are new and 10 of which were presented one time in the original series.

Recognition of Faces-Delayed Non-Match to Sample: On each of the 24 trials, faces are added to an array, with the subject being required to identify the new face added by touching it on the monitor.

Telephone Dialing (with and without interference): This task is a variation on the standard digit repetition paradigm. Participants are

instructed to dial long distance telephone numbers on from memory on a touchtone phone interfaced with the computer. The interference condition is identical to the no interference condition, with the exception that after the subjects have completed dialing they hear a busy signal and they are asked to redial.

First-Last Names: This is a verbal learning test using the WMS paired associates paradigm. Subjects are presented with a series of six first-last name pairs, followed by a presentation of the last name only, with subjects recalling the first. Total name-face pairs recall are summed across five learning trials.

Grocery List Selective Reminding Test: This test follows the standard selective reminding paradigm (Buschke, 1973) using common grocery-list items as the stimuli to be recalled, presented on the video screen. After 40 minutes, delayed recall is also assessed.

Traditional Tests: The following traditional instruments were administered in the standard format: Logical Memory and Paired Associate Learning from the Wechsler Memory Scale (WMS; Wechsler, 1945), the Digit Span and Digit Symbol subtests from the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955), and the Benton Visual Retention Test (BVRT; Benton, 1974).

RESULTS

Inspection of Table 1 reveals that all of the scores yielded by these tests were significantly correlated for the two administrations at the .01 alpha level, with the exception of Recognition of Faces-Signal Detection-40 minute delay D' ($r = .18$). Those correlations that were significant ranged from a high of .83 for Name-Face Association-Total Recall to a low of .32 for Recognition of Faces-Delayed Non-Matching-to-Sample-First Miss. The magnitude of the practice effects was examined using univariate repeated measures

analyses of variance. The resulting alpha levels for each of the tests in the MAC battery are presented in Table 1.

Table 2 presents the means, standard deviations, reliability coefficients with significance levels, and the alpha levels of the practice effects for the five traditional memory/attention tests. All test-retest correlations were significant at the .01 alpha level. They ranged from a high of .88 for WAIS Digit Symbol to a low of .53 for Benton Visual Retention Test-Number of Errors. Inspection of Table 2 also reveals significant practice effects for all traditional test scores, with the exception of WAIS Digit Span, which was marginally significant.

DISCUSSION

The test-retest reliability of the recall memory measures from the MAC battery generally compared favorably to the traditional memory tests. With respect to those measures that are similar in structure, the computerized, everyday measure of prose recall, Divided Attention Recall, had a greater test-retest correlation than its counterpart, WMS Logical Memory. The test-retest correlations for First-Last Names, a computerized, everyday verbal paired associate learning task, and WMS Paired Associate Learning were roughly equivalent. Name-Face Association, another paired associate measure, had higher test-retest reliability than either First-Last Names or WMS Paired Associate Learning. The other computerized, everyday verbal learning task, the Grocery List Selective Reminding Test, had better test-retest reliability than Logical Memory and was roughly equivalent to Paired Associate Learning.

With respect to visual memory measures, the test-retest stability of Misplaced Objects was superior to the Benton Visual Retention Test (BVRT), Recognition of Faces-Delayed Non-Match-to-Sample-Total Correct was roughly equivalent to the BVRT, and the test-retest reliability of the BVRT was

superior to most scores from Recognition of Faces-Signal Detection.

It was apparent that the traditional psychometric measures of attention and concentration, i.e., WAIS Digit Span and Digit Symbol, had superior test-retest reliability to the computerized attention and concentration measures. Specifically, of the two measures of serial digit recall, Digit Span had better test-rest stability than Telephone Dialing. The best test-retest reliability of all measures (computerized and standard) was obtained with WAIS Digit Symbol.

Table 1

Test-Retest Reliability and Practice Effects of Computerized, Everyday Memory Tests

	Visit 1 Mean (SD)	Visit 2 Mean (SD)	Reliability Coefficients	Practice Effect Alpha Levels
Simple Reaction Time				
Lift Time	218.71 (48.11)	231.40 (70.04)	.47*	.034
Travel Time	232.72 (132.71)	222.45 (117.44)	.44*	.409
Divided Attention				
Lift Time	324.77 (112.22)	330.62 (105.52)	.49*	.575
Travel Time	247.87 (101.94)	242.39 (99.77)	.68*	.472
Averaged Recall	21.64 (8.54)	27.94 (11.84)	.72*	.000
Misplaced Objects				
Found on First Attempt	12.40 (4.05)	12.63 (4.50)	.79*	.388
Total Found	15.05 (3.23)	15.33 (3.67)	.78*	.199

(table continues)

	Visit 1 Mean (SD)		Visit 2 Mean (SD)		Reliability Coefficients	Practice Effect Alpha Levels
Name-Face Associations						
Total Recall Over 3 Learning Trials	24.92	(9.08)	34.09	(8.71)	.83*	.000
Delayed Recall	10.82	(3.61)	12.47	(2.95)	.80*	.000
Recognition of Faces - Signal Detection						
Immediate D'	2.63	(.82)	2.87	(.89)	.60*	.001
One minute D'	2.06	(.65)	2.07	(.76)	.37*	.867
Two minute D'	1.88	(.69)	2.00	(.68)	.50*	.060
Three minute D'	1.76	(.66)	1.94	(.71)	.37*	.010
Four minute D'	1.28	(.68)	1.54	(.70)	.34*	.001
40 minute D'	.42	(.52)	.46	(.50)	.18	.622
Recognition of Faces - Delayed Non-Match-to-Sample						
First Miss	11.59	(5.77)	12.49	(6.15)	.32*	.181
Total Correct	18.03	(4.38)	18.84	(4.05)	.56*	.030

(table continues)

	Visit 1 Mean (SD)		Visit 2 Mean (SD)		Reliability Coefficients	Practice Effect Alpha Levels
Telephone Dialing						
10 Digits without Interference	7.13	(1.47)	7.48	(1.48)	.48*	.021
10 Digits After Interference	4.71	(2.20)	5.00	(2.36)	.54*	.185
First-Last Names						
Total Recall Over 5 Learning Trials	19.71	(6.58)	24.10	(5.72)	.69*	.000
Grocery List Selective Reminding						
Total Recall Over 5 Learning Trials	58.25	(9.88)	67.20	(9.14)	.74*	.000
Delayed Recall	12.46	(2.48)	13.77	(1.94)	.64*	.000

* $p < .01$

Table 2

Test-Retest Reliability and Practice Effects of Selected Traditional Non-Computerized Tests

	Visit 1 Mean (SD)		Visit 2 Mean (SD)		Reliability Coefficients	Practice Effect Alpha Levels
Logical Memory	8.74	(2.33)	10.17	(2.56)	.55*	.000
Paired Associate Learning	16.59	(3.70)	17.90	(3.48)	.72*	.000
Benton Visual Retention Test						
Number Correct	7.05	(1.54)	7.65	(1.62)	.57*	.000
Number Errors	3.76	(2.13)	3.03	(2.43)	.53*	.001
Digit Span	13.01	(2.10)	13.30	(2.14)	.73*	.045
Digit Symbol	58.75	(13.60)	61.63	(14.15)	.88*	.000

* $p < .01$